

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Reed, John
- (ii) TITLE OF INVENTION: Regulation of bcl-2 Gene Expression
- (iii) NUMBER OF SEQUENCES: 23
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Gray, Cary, Ames & Frye
 - (B) STREET: 401 B Street, Suite 1700
 - (C) CITY: San Diego
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 92101-4297
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb storage
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: US
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Brotman, Harris F.
 - (B) REGISTRATION NUMBER: 35461
 - (C) REFERENCE/DOCKET NUMBER: P0041US0
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (619) 699-3630
 - (B) TELEFAX: (619) 236-1048

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iv) ANTI-SENSE: YES
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCCTTCTAC CGCGTGCGAC

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CTTTCCTCT GGAAGGATG GCGCACGCTG GGAGA

35

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCTCCGACCC ATCCACGTAG

20

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 33 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACGGGGTACG GAGGCTGGGT AGGTGCATCT GGT

33

662700455260

*Sub
C-1
cont*

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTTGACGTCC TACGGAAACA

20

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCCCCAACTG CAGGATGCCT TTGTGGA ACT GTACGG

36

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGGAAGGATG GCGCACGCTG

20

(2) INFORMATION FOR SEQ ID NO:8:

682780-1155260

*See
C1
cont.*

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGTGCGAC CCTCTTG

17

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TACCGCGTGC GACCCTC

17

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TCCTACCGCG TGCGACC

17

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid

002544-0347200

*Sub
C1
cmi*

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

17

CCTTCCTACC GCGTGCG

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

17

GACCCTTCCT ACCGCGT

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

17

GGAGACCCTT CCTACCG

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

662700-1755260

Sub
C1
Cent

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

15

CGGGCGGCAG CGCGG

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

15

CGGGCGGGCG ACGGA

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

16

CGGGAGCGCG GCGGGC

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

CGGGAGCGCG GCGGGC

*Sub
C1
Cont.*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TCTCCCAGCG TGCGGCAT

18

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGCACTCACG CTCGGCCT

18

*Sub
C1
cont.*

002700"44554E00

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5086 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCGCCCGCCC CTCCGCGCCG CCTGCCCCGCC CGCCCGCCGC GCTCCCGCCC GCCGCTCTCC	60
GTGGCCCCGC CGCGCTGCCG CCGCCGCCGC TGCCAGCGAA GGTGCCGGGG CTCCGGGCCC	120
TCCCTGCCGG CGGCCGTCAG CGCTCGGAGC GAACTGCGCG ACGGGAGGTC CGGGAGGCGA	180
CCGTAGTCGC GCCGCCGCGC AGGACCAGGA GGAGGAGAAA GGGTGCGCAG CCCGGAGGCG	240
GGGTGCGCCG GTGGGGTGCA GCGGAAGAGG GGGTCCAGGG GGGAGAACTT CGTAGCAGTC	300
ATCCTTTTTTA GGAAAAGAGG GAAAAAATAA AACCTCCCC CACCACCTCC TTCTCCCCAC	360
CCCTCGCCGC ACCACACACA GCGCGGGCTT CTAGCGCTCG GCACCGGCGG GCCAGGCGCG	420
TCCTGCCTTC ATTTATCCAG CAGCTTTTCG GAAAATGCAT TTGCTGTTTC GAGTTTAATC	480
AGAAGACGAT TCCTGCCTCC GTCCCGGGCT CCTTCATCGT CCCATCTCCC CTGTCTCTCT	540
CCTGGGGAGG CGTGAAGCGG TCCCGTGGAT AGAGATTCAT GCCTGTGTCC GCGCGTGTGT	600
GCGCGCGTAT AAATTGCCGA GAAGGGGAAA ACATCACAGG ACTTCTGCGA ATACCGGACT	660
GAAAATTGTA ATTCATCTGC CGCCGCCGCT GCCAAAAAAA AACTCGAGCT CTTGAGATCT	720
CCGGTTGGGA TTCCTGCGGA TTGACATTTT TGTGAAGCAG AAGTCTGGGA ATCGATCTGG	780
AAATCCTCCT AATTTTTACT CCCTCTCCCC CCGACTCCTG ATTCATTGGG AAGTTTCAA	840
TCAGCTATAA CTGGAGAGTG CTGAAGATTG ATGGGATCGT TGCCTTATGC ATTTGTTTTG	900
GTTTTACAAA AAGGAAACTT GACAGAGGAT CATGCTGTAC TTA AAAAATA CAAGTAAGTC	960
TCGCACAGGA AATTGGTTTA ATGTAACCTT CAATGGAAAC CTTTGAGATT TTTTACTTAA	1020
AGTGCAATCG AGTAAATTTA ATTTCCAGGC AGCTTAATAC ATTGTTTTTA GCCGTGTTAC	1080
TTGTAGTGTG TATGCCCTGC TTTCACTCAG TGTGTACAGG GAAACGCACC TGATTTTTTA	1140
CTTATTAGTT TGTTTTTTCT TTAACCTTTC AGCATCACAG AGGAAGTAGA CTGATATTAA	1200
CAATACTTAC TAATAATAAC GTGCCTCATG AAATAAAGAT CCGAAAGGAA TTGGAATAAA	1260
AATTCCTGCT GTCTCATGCC AAGAGGGAAA CACCAGAATC AAGTGTTCCG CGTGATTGAA	1320

664730 1459260

*Sub
C1
cont*

GACACCCCCT	CGTCCAAGAA	TGCAAAGCAC	ATCCAATAAA	ATAGCTGGAT	TATAACTCCT	1380
CTTCTTTTCTC	TGGGGGCCGT	GGGGTGGGAG	CTGGGGCGAG	AGGTGCCGTT	GGCCCCCGTT	1440
GCTTTTCCTC	TGGGAAGGAT	GGCGCACGCT	GGGAGAACGG	GGTACGACAA	CCGGGAGATA	1500
GTGATGAAGT	ACATCCATTA	TAAGCTGTCT	CAGAGGGGCT	ACGAGTGGGA	TGCGGGAGAT	1560
GTGGGCGCCG	CGCCCCCGGG	GGCCGCCCCC	GCACCGGGCA	TCTTCTCCTC	CCAGCCCGGG	1620
CACACGCCCC	ATCCAGCCGC	ATCCCGCGAC	CCGGTCGCCA	GGACCTCGCC	GCTGCAGACC	1680
CCGGCTGCCC	CCGGCGCCGC	CGCGGGGCCT	GCGCTCAGCC	CGGTGCCACC	TGTGGTCCAC	1740
CTGGCCCTCC	GCCAAGCCGG	CGACGACTTC	TCCCGCCGCT	ACCGCGGCGA	CTTCGCCGAG	1800
ATGTCCAGCC	AGCTGCACCT	GACGCCCTTC	ACCGCGCGGG	GACGCTTTGC	CACGGTGGTG	1860
GAGGAGCTCT	TCAGGGACGG	GGTGAAGTGG	GGGAGGATTG	TGGCCTTCTT	TGAGTTCGGT	1920
GGGGTCATGT	GTGTGGAGAG	CGTCAACCGG	GAGATGTCGC	CCCTGGTGGA	CAACATCGCC	1980
CTGTGGATGA	CTGAGTACCT	GAACCGGCAC	CTGCACACCT	GGATCCAGGA	TAACGGAGGC	2040
TGGGATGCCT	TTGTGGAAGT	GTACGGCCCC	AGCATGCGGC	CTCTGTTTGA	TTTCTCCTGG	2100
CTGTCTCTGA	AGACTCTGCT	CAGTTTGGCC	CTGGTGGGAG	CTTGCAACAC	CCTGGGTGCC	2160
TATCTGAGCC	ACAAGTGAAG	TCAACATGCC	TGCCCCAAAC	AAATATGCAA	AAGGTTCACT	2220
AAAGCAGTAG	AAATAATATG	CATTGTCACT	GATGTACCAT	GAAACAAAGC	TGCAGGCTGT	2280
TTAAGAAAAA	ATAACACACA	TATAAACATC	ACACACACAG	ACAGACACAC	ACACACACAA	2340
CAATTAACAG	TCTTCAGGCA	AAACGTCGAA	TCAGCTATTT	ACTGCCAAAG	GGAAATATCA	2400
TTTATTTTTT	ACATTATTAA	GAAAAAAGAT	TTATTTATTT	AAGACAGTCC	CATCAAAACT	2460
CCGTCTTTGG	AAATCCGACC	ACTAATTGCC	AAACACCGCT	TCGTGTGGCT	CCACCTGGAT	2520
GTTCTGTGCC	TGTAAACATA	GATTCGCTTT	CCATGTTGTT	GGCCGGATCA	CCATCTGAAG	2580
AGCAGACGGA	TGGAAAAAGG	ACCTGATCAT	TGGGGAAGCT	GGCTTTCTGG	CTGCTGGAGG	2640
CTGGGGAGAA	GGTGTTCATT	CACTTGCATT	TCTTTGCCCT	GGGGGCGTGA	TATTAACAGA	2700
GGGAGGGTTC	CCGTGGGGGG	AAGTCCATGC	CTCCCTGGCC	TGAAGAAGAG	ACTCTTTGCA	2760
TATGACTCAC	ATGATGCATA	CCTGGTGGGA	GGAAAAGAGT	TGGGAAGTTC	AGATGGACCT	2820
AGTACCCACT	GAGATTTCCA	CGCCGAAGGA	CAGCGATGGG	AAAAATGCCC	TTAAATCATA	2880
GGAAAGTATT	TTTTTAAGCT	ACCAATTGTG	CCGAGAAAAG	CATTTTAGCA	ATTTATACAA	2940
TATCATCCAG	TACCTTAAAC	CCTGATTGTG	TATATTCATA	TATTTTGGAT	ACGCACCCCC	3000

664753" 47532660

Sub
C-
cont.

CAACTCCCAA	TACTGGCTCT	GTCTGAGTAA	GAAACAGAAT	CCTCTGGAAC	TTGAGGAAGT	3060
GAACATTTTCG	GTGACTTCCG	ATCAGGAAGG	CTAGAGTTAC	CCAGAGCATC	AGGCCGCCAC	3120
AAGTGCCTGC	TTTTAGGAGA	CCGAAGTCCG	CAGAACCTAC	CTGTGTCCCA	GCTTGGAGGC	3180
CTGGTCCTGG	AACTGAGCCG	GGCCCTCACT	GGCCTCCTCC	AGGGATGATC	AACAGGGTAG	3240
TGTGGTCTCC	GAATGTCTGG	AAGCTGATGG	ATGGAGCTCA	GAATTCCACT	GTCAAGAAAG	3300
AGCAGTAGAG	GGGTGTGGCT	GGGCCTGTCA	CCCTGGGGCC	CTCCAGGTAG	GCCCCTTTTC	3360
ACGTGGAGCA	TAGGAGCCAC	GACCCTTCTT	AAGACATGTA	TCACTGTAGA	GGGAAGGAAC	3420
AGAGGCCCTG	GGCCTTCCTA	TCAGAAGGAC	ATGGTGAAGG	CTGGGAACGT	GAGGAGAGGC	3480
AATGGCCACG	GCCCATTTTG	GCTGTAGCAC	ATGGCACGTT	GGCTGTGTGG	CCTTGGCCAC	3540
CTGTGAGTTT	AAAGCAAGGC	TTTAAATGAC	TTTGGAGAGG	GTCACAAATC	CTAAAAGAAG	3600
CATTGAAGTG	AGGTGTCATG	GATTAAATTGA	CCCCTGTCTA	TGGAATTACA	TGTAAAACAT	3660
TATCTTGTC	CTGTAGTTTG	GTTTTATTG	AAAACCTGAC	AAAAAAAAG	TTCCAGGTGT	3720
GGAATATGGG	GGTTATCTGT	ACATCCTGGG	GCATTAAAAA	AAAATCAATG	GTGGGGAAC	3780
ATAAAGAAGT	AACAAAAGAA	GTGACATCTT	CAGCAAATAA	ACTAGGAAAT	TTTTTTTTCT	3840
TCCAGTTTAG	AATCAGCCTT	GAAACATTGA	TGGAATAACT	CTGTGGCATT	ATTGCATTAT	3900
ATACCATTTA	TCTGTATTAA	CTTTGGAATG	TACTCTGTTC	AATGTTTAAT	GCTGTGGTTG	3960
ATATTTGAA	AGCTGCTTTA	AAAAAATACA	TGCATCTCAG	CGTTTTTTTG	TTTTTAATTG	4020
TATTTAGTTA	TGGCCTATAC	ACTATTTGTG	AGCAAAGGTG	ATCGTTTTCT	GTTTGAGATT	4080
TTTATCTCTT	GATTCTTCAA	AAGCATTCTG	AGAAGGTGAG	ATAAGCCCTG	AGTCTCAGCT	4140
ACCTAAGAAA	AACCTGGATG	TCACTGGCCA	CTGAGGAGCT	TTGTTTCAAC	CAAGTCATGT	4200
GCATTTCCAC	GTCAACAGAA	TTGTTTATTG	TGACAGTTAT	ATCTGTTGTC	CCTTTGACCT	4260
TGTTTCTTGA	AGGTTTCCTC	GTCCCTGGGC	AATCCGCAT	TTAATTCATG	GTATTCAGGA	4320
TTACATGCAT	GTTTGTTTAA	ACCCATGAGA	TTCATTCAGT	TAAAAATCCA	GATGGCGAAT	4380
GACCAGCAGA	TTCAAATCTA	TGGTGGTTTG	ACCTTTAGAG	AGTTGCTTTA	CGTGGCCTGT	4440
TTCAACACAG	ACCCACCCAG	AGCCCTCCTG	CCCTCCTTCC	GCGGGGGCTT	TCTCATGGCT	4500
GTCCTTCAGG	GTCTTCCTGA	AATGCAGTGG	TCGTTACGCT	CCACCAAGAA	AGCAGGAAAC	4560
CTGTGGTATG	AAGCCAGACC	TCCCCGGCGG	GCCTCAGGGA	ACAGAATGAT	CAGACCTTTG	4620
AATGATTCTA	ATTTTAAAGC	AAAATATTAT	TTTATGAAAG	GTTTACATTG	TCAAAGTGAT	4680

664700" 4554660

Sub
C1
cont

GAATATGGAA TATCCAATCC TGTGCTGCTA TCCTGCCAAA ATCATTTTAA TGGAGTCAGT 4740
TTGCAGTATG CTCCACCTGG TAAGATCCTC CAAGCTGCTT TAGAAGTAAC AATGAAGAAC 4800
GTGGACGTTT TTAATATAAA GCCTGTTTTG TCTTTTGTTG TTGTTCAAAC GGGATTCACA 4860
GAGTATTTGA AAAATGTATA TATATTAAGA GGTCACGGGG GCTAATTGCT AGCTGGCTGC 4920
CTTTTGCTGT GGGGTTTTGT TACCTGGTTT TAATAACAGT AAATGTGCCC AGCCTCTTGG 4980
CCCCAGAACT GTACAGTATT GTGGCTGCAC TTGCTCTAAG AGTAGTTGAT GTTGCATTTT 5040
CCTTATTGTT AAAACATGT TAGAAGCAAT GAATGTATAT AAAAGC 5086

*Sub
cl
cont*

002544.084700

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 717 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..717

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATG GCG CAC GCT GGG AGA ACG GGG TAC GAC AAC CGG GAG ATA GTG ATG	48
Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met	
1 5 10 15	
AAG TAC ATC CAT TAT AAG CTG TCG CAG AGG GGC TAC GAG TGG GAT GCG	96
Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala	
20 25 30	
GGA GAT GTG GGC GCC GCG CCC CCG GGG GCC GCC CCC GCA CCG GGC ATC	144
Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile	
35 40 45	
TTC TCC TCC CAG CCC GGG CAC ACG CCC CAT CCA GCC GCA TCC CGC GAC	192
Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp	
50 55 60	
CCG GTC GCC AGG ACC TCG CCG CTG CAG ACC CCG GCT GCC CCC GGC GCC	240
Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala	
65 70 75 80	
GCC GCG GGG CCT GCG CTC AGC CCG GTG CCA CCT GTG GTC CAC CTG GCC	288
Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala	
85 90 95	
CTC CGC CAA GCC GGC GAC GAC TTC TCC CGC CGC TAC CGC GGC GAC TTC	336
Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe	
100 105 110	
GCC GAG ATG TCC AGC CAG CTG CAC CTG ACG CCC TTC ACC GCG CGG GGA	384
Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly	
115 120 125	
CGC TTT GCC ACG GTG GTG GAG GAG CTC TTC AGG GAC GGG GTG AAC TGG	432
Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp	
130 135 140	

662160" 4452260

*Sub
C1
cont*

GGG	AGG	ATT	GTG	GCC	TTC	TTT	GAG	TTC	GGT	GGG	GTC	ATG	TGT	GTG	GAG	480
Gly	Arg	Ile	Val	Ala	Phe	Phe	Glu	Phe	Gly	Gly	Val	Met	Cys	Val	Glu	
145					150					155					160	
AGC	GTC	AAC	CGG	GAG	ATG	TCG	CCC	CTG	GTG	GAC	AAC	ATC	GCC	CTG	TGG	528
Ser	Val	Asn	Arg	Glu	Met	Ser	Pro	Leu	Val	Asp	Asn	Ile	Ala	Leu	Trp	
				165					170					175		
ATG	ACT	GAG	TAC	CTG	AAC	CGG	CAC	CTG	CAC	ACC	TGG	ATC	CAG	GAT	AAC	576
Met	Thr	Glu	Tyr	Leu	Asn	Arg	His	Leu	His	Thr	Trp	Ile	Gln	Asp	Asn	
			180					185					190			
GGA	GGC	TGG	GAT	GCC	TTT	GTG	GAA	CTG	TAC	GGC	CCC	AGC	ATG	CGG	CCT	624
Gly	Gly	Trp	Asp	Ala	Phe	Val	Glu	Leu	Tyr	Gly	Pro	Ser	Met	Arg	Pro	
		195					200					205				
CTG	TTT	GAT	TTC	TCC	TGG	CTG	TCT	CTG	AAG	ACT	CTG	CTC	AGT	TTG	GCC	672
Leu	Phe	Asp	Phe	Ser	Trp	Leu	Ser	Leu	Lys	Thr	Leu	Leu	Ser	Leu	Ala	
	210					215					220					
CTG	GTG	GGA	GCT	TGC	ATC	ACC	CTG	GGT	GCC	TAT	CTG	AGC	CAC	AAG		717
Leu	Val	Gly	Ala	Cys	Ile	Thr	Leu	Gly	Ala	Tyr	Leu	Ser	His	Lys		
225					230					235						

*Sub
C1
cont*

003554 09199

(A) LENGTH: 239 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
1 5 10 15
Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
20 25 30
Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
35 40 45
Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
50 55 60
Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
65 70 75 80
Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala
85 90 95
Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe
100 105 110
Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
115 120 125
Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp
130 135 140
Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145 150 155 160
Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
165 170 175
Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
180 185 190
Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly Pro Ser Met Arg Pro
195 200 205
Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr Leu Leu Ser Leu Ala
210 215 220

Sub
C1
cont

Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr Leu
225 230 235

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 615 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..615

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATG GCG CAC GCT GGG AGA ACG GGG TAC GAC AAC CGC
Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg
1 5 10

AAG TAC ATC CAT TAT AAG CTG TCG CAG AGG GGC TAC
Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr
20 25

GGA GAT GTG GGC GCC GCG CCC CCG GGG GCC GCC CCC
Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro
35 40

TTC TCC TCC CAG CCC GGG CAC ACG CCC CAT CCA GCC
Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala
50 55 60

CCG GTC GCC AGG ACC TCG CCG CTG CAG ACC CCG GCT
Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala
65 70 75

GCC GCG GGG CCT GCG CTC AGC CCG GTG CCA CCT GTG
Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val
85 90

- ```
(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..615
```

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATG | GCG | CAC | GCT | GGG | AGA | ACG | GGG | TAC | GAC | AAC | CGG | GAG | ATA | GTG | ATG | 48  |
| Met | Ala | His | Ala | Gly | Arg | Thr | Gly | Tyr | Asp | Asn | Arg | Glu | Ile | Val | Met |     |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |     |
| AAG | TAC | ATC | CAT | TAT | AAG | CTG | TCG | CAG | AGG | GGC | TAC | GAG | TGG | GAT | GCG | 96  |
| Lys | Tyr | Ile | His | Tyr | Lys | Leu | Ser | Gln | Arg | Gly | Tyr | Glu | Trp | Asp | Ala |     |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |     |
| GGA | GAT | GTG | GGC | GCC | GCG | CCC | CCG | GGG | GCC | GCC | CCC | GCA | CCG | GGC | ATC | 144 |
| Gly | Asp | Val | Gly | Ala | Ala | Pro | Pro | Gly | Ala | Ala | Pro | Ala | Pro | Gly | Ile |     |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |     |
| TTC | TCC | TCC | CAG | CCC | GGG | CAC | ACG | CCC | CAT | CCA | GCC | GCA | TCC | CGC | GAC | 192 |
| Phe | Ser | Ser | Gln | Pro | Gly | His | Thr | Pro | His | Pro | Ala | Ala | Ser | Arg | Asp |     |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |     |
| CCG | GTC | GCC | AGG | ACC | TCG | CCG | CTG | CAG | ACC | CCG | GCT | GCC | CCC | GGC | GCC | 240 |
| Pro | Val | Ala | Arg | Thr | Ser | Pro | Leu | Gln | Thr | Pro | Ala | Ala | Pro | Gly | Ala |     |
| 65  |     |     |     |     | 70  |     |     |     | 75  |     |     |     |     |     | 80  |     |
| GCC | GCG | GGG | CCT | GCG | CTC | AGC | CCG | GTG | CCA | CCT | GTG | GTC | CAC | CTG | GCC | 288 |
| Ala | Ala | Gly | Pro | Ala | Leu | Ser | Pro | Val | Pro | Pro | Val | Val | His | Leu | Ala |     |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |     |

430 434 436 438 440 442 444 446 448 450 452 454 456 458 460 462 464 466 468 470 472 474 476 478 480 482 484 486 488 490 492 494 496 498 500 502 504 506 508 510 512 514 516 518 520 522 524 526 528 530 532 534 536 538 540 542 544 546 548 550 552 554 556 558 560 562 564 566 568 570 572 574 576 578 580 582 584 586 588 590 592 594 596 598 600 602 604 606 608 610 612 614 616 618 620 622 624 626 628 630 632 634 636 638 640 642 644 646 648 650 652 654 656 658 660 662 664 666 668 670 672 674 676 678 680 682 684 686 688 690 692 694 696 698 700 702 704 706 708 710 712 714 716 718 720 722 724 726 728 730 732 734 736 738 740 742 744 746 748 750 752 754 756 758 760 762 764 766 768 770 772 774 776 778 780 782 784 786 788 790 792 794 796 798 800 802 804 806 808 810 812 814 816 818 820 822 824 826 828 830 832 834 836 838 840 842 844 846 848 850 852 854 856 858 860 862 864 866 868 870 872 874 876 878 880 882 884 886 888 890 892 894 896 898 900 902 904 906 908 910 912 914 916 918 920 922 924 926 928 930 932 934 936 938 940 942 944 946 948 950 952 954 956 958 960 962 964 966 968 970 972 974 976 978 980 982 984 986 988 990 992 994 996 998 1000

GGA GGC TGG GTA GGT GCA TCT ~~GGT~~ GAT GTG AGT CTG GGC  
Gly Gly Trp Val Gly Ala Ser ~~Gly~~ Asp Val Ser Leu Gly 615  
195 200 205

Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala  
65 70 75 80

Sub  
C1  
cont.



|     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| Pro | Ala | Leu | Ser | Pro | Val | 85  |
| Ala | Gly | Asp | Asp | Phe | Ser | 100 |
| Ser | Ser | Gln | Leu | His | Leu | 120 |
| Thr | Val | Val | Glu | Glu | Leu | 135 |
| Val | Ala | Phe | Phe | Glu | Phe | 150 |
| Arg | Glu | Met | Ser | Pro | Leu | 165 |
| Tyr | Leu | Asn | Arg | His | Leu | 180 |
| Val | Gly | Ala | Ser | Gly | Asp | 200 |

**SECRET**